# On Aliased Resizing and Surprising Subtleties in GAN Evaluation

Year: 2022 | Citations: 450 | Authors: Gaurav Parmar, Richard Zhang, Jun-Yan Zhu

---

## Abstract

Metrics for evaluating generative models aim to measure the discrepancy between real and generated images. The often-used Fréchet Inception Distance (FID) metric, for example, extracts "high-level" features using a deep network from the two sets. However, we find that the differences in "low-level" preprocessing, specifically image resizing and compression, can induce large variations and have unforeseen consequences. For instance, when resizing an image, e.g., with a bilinear or bicubic kernel, signal processing principles mandate adjusting prefilter width depending on the downsampling factor, to antialias to the appropriate bandwidth. However, commonly-used implementations use a fixed-width prefilter, resulting in aliasing artifacts. Such aliasing leads to corruptions in the feature extraction down-stream. Next, lossy compression, such as JPEG, is commonly used to reduce the file size of an image. Although designed to minimally degrade the perceptual quality of an image, the operation also produces variations downstream. Furthermore, we show that if compression is used on real training images, FID can actually improve if the generated images are also subsequently compressed. This paper shows that choices in low-level image processing have been an under-appreciated aspect of generative modeling. We identify and characterize variations in generative modeling development pipelines, provide recommendations based on signal processing principles, and release a reference implementation to facilitate future comparisons.